Annals of Operations Research 81(1998)3–14

A nonconvex weighted potential function for polynomial target following methods

E. de Klerk, C. Roos and T. Terlaky

Faculty of Technical Mathematics and Informatics,
Delft University of Technology, P.O. Box 5031,
2600 GA Delft, The Netherlands

E-mail: {e.deklerk; c.roos; t.terlaky}@twi.tudelft.nl

Long step interior-point methods in linear programming are some of the most efficient algorithms from a computational point of view. We prove polynomial complexity of a class of long step target-following methods in a novel way, by introducing a new nonconvex potential function and adapting the analysis framework of Jansen et al. [4, 6, 7]. The main advantage is that the new potential function has an obvious extension to semi-definite programming, whereas the potential used in the above-mentioned papers does not.

Keywords: interior-point method, primal-dual method, target-following, Dikin steps

1. Introduction

Medium and long step primal-dual interior-point methods in linear programming are of significant practical importance. Introduced by Kojima et al. [9], these methods have proven efficient in computational studies [11].

The worst-case complexity of long step algorithms with $O(1)$ barrier parameter (target) updates is $O(n \ln 1/\varepsilon)$ iterations, and for medium updates of $O(1/\sqrt{n})$, one has a worst-case bound of $O(\sqrt{n} \ln 1/\varepsilon)$ iterations [2, 3, 8]. Although the long step methods have a worse complexity bound than the short and medium step variants, the number of iterations performed in practice are often lower, as becomes clear from the cited references.

Jansen et al. [4, 6, 7] provided a unifying framework of analysis for these important algorithms. Their “target-following” approach involves choosing a series of targets to be approximated in the primal-dual space.

To introduce this approach, consider an LP problem in standard form

$$\min_x \{c^T x : Ax = b,\ x \in \mathbb{R}^n_+\},$$

where $\mathbb{R}^n_+$ denotes the positive orthant in $\mathbb{R}^n$, and its dual problem

$$\max_{y,s}\{b^T y : A^T y + s = c,\ s \in \mathbb{R}^n_+, y \in \mathbb{R}^m\}.$$

For each target in the positive orthant, say $\bar{v} \in \mathbb{R}^n_+$, there exists a unique primal-dual feasible pair (x, s) such that[1] $xs = \bar{v}^2$. Since all optimal pairs satisfy $xs = 0$, it is natural to choose a sequence of targets $\{\bar{v}^{(j)}\}$ in the positive orthant which converges to zero, and to compute a pair $(x^{(j)}, s^{(j)})$ such that $x^{(j)}s^{(j)} \approx (\bar{v}^{(j)})^2$ for each target in the sequence $\bar{v}^{(0)}, \bar{v}^{(1)}, \ldots$. Denoting $v^2 = xs$ for any primal-dual pair (x, s), we can make the approximation relation "$\approx$" more precise by using the proximity measure introduced by Jansen et al. in [6]:

$$\delta(v, \bar{v}) = \frac{1}{2\min(\bar{v})}\left\|\frac{\bar{v}^2 - v^2}{v}\right\|,$$

where $\min(\bar{v}) := \min_{1 \le i \le n}\{\bar{v}_i\}$. We say v^2 is close to $\bar{v}^2$ if $\delta(v, \bar{v}) \le \tau$ for some tolerance $\tau < 1$.

The pair $(x^{(j)}, s^{(j)})$ is obtained by (approximately) solving the nonlinear system

$$\begin{aligned} Ax &= b, & x &\ge 0, \\ A^T y + s &= c, & s &\ge 0, \\ xs &= (\bar{v}^{(j)})^2. \end{aligned}$$

This is done iteratively by a damped Newton method, i.e. by taking damped Newton steps until the approximation condition is satisfied. The pairs $(x^{(j)}, s^{(j)})$ are called outer iterates, and the points generated during the Newton process will be termed inner iterates.

The Newton step $(\Delta x, \Delta s)$ is obtained by solving the linearized system

$$\left.\begin{aligned} A\Delta x &= 0, \\ A^T\Delta y + \Delta s &= 0, \\ x\Delta s + s\Delta x &= (\bar{v}^{(j)})^2 - v^2, \end{aligned}\right\} \tag{1}$$

where the pair (x, s) is the last pair of inner iterates.

A damped Newton step $(\alpha\Delta x, \alpha\Delta s)$ with $\alpha \le 1$ is used (as opposed to a full Newton step) and some care is required in choosing the step length α to ensure convergence of Newton's method.

To this end, a potential function is used in the analysis of the Newton process. The idea is that a sufficient reduction in the potential ensures proximity of the Newton iterates to the target $\bar{v}^{(j)}$. The analysis therefore reduces to analysing the effect of the damped Newton steps on the potential. (In practice the potential function may be used

[1] We use componentwise notation: xs indicates the vector obtained by multiplying the corresponding components of x and s, v^2 is the vector obtained by squaring the components of v, etc.

in line searches to do larger steps than allowed for by the analysis.) It is shown that a step length $\alpha \leq 1$ may be found at each step, which ensures a decrease of the potential by an absolute constant.

The target $\overline{v}^{(j)}$ is updated as soon as the proximity condition is satisfied, i.e. as soon as the potential has been sufficiently decreased.

The result is the conceptually appealing *target-following framework*:

Target-following algorithm

Initialization

Given an initial feasible pair $(x^{(0)}, s^{(0)})$;
Let $\varepsilon > 0$ be an accuracy parameter and $\tau < 1$ a proximity parameter;
Choose an initial target $\overline{v}^{(0)}$ such that $\delta(v^{(0)}, \overline{v}^{(0)}) \leq \tau$.
Set counter $j = 0$, $x = x^{(0)}$, and $s = s^{(0)}$.

While $(x^{(j)})^T s^{(j)} > \varepsilon$ **do**

(1) Solve the Newton equations (1) to obtain Δx and Δs.

(2) Choose a suitable damping parameter (step length) $\alpha \leq 1$.

(3) Set $x = x + \alpha \Delta x$, $s = s + \alpha \Delta s$, $v = \sqrt{xs}$.

(4) If $\delta(v, \overline{v}^{(j)}) \leq \tau$, then
- Let $(x^{(j+1)}, s^{(j+1)}) = (x, s)$;
- Choose a new target $\overline{v}^{(j+1)}$;
- Set $j = j + 1$.

Enddo

The primal-dual potential function used in the papers [4, 6, 7] to determine the step length α is a strictly convex function. A new potential function is introduced here which is nonconvex but still suitable for the complexity analysis of long step algorithms. The advantage of the new function is that it has an obvious analogue in the semi-definite programming case, whereas the potential used in [4, 6, 7] does not.

2. A new potential function

The new potential function used here is

$$f(x, s, \overline{v}) = \sum_{i=1}^{n} (x_i s_i \overline{v}_i^{-2} - 1 - \ln x_i s_i \overline{v}_i^{-2}), \tag{2}$$

defined on the primal-dual feasible region. Using $v = \sqrt{xs}$, we can write (2) as

$$\phi(v, \overline{v}) = \sum_{i=1}^{n} \left(\frac{v_i^2}{\overline{v}_i^2} - 1 - \ln \frac{v_i^2}{\overline{v}_i^2} \right). \tag{3}$$

Note that $\phi(v, \overline{v}) \geq \phi(\overline{v}, \overline{v}) = 0$.

The proposed potential function differs from the potential used by Jansen et al. in [4, 6, 7],

$$\tilde{f}(x, s, \bar{v}) = \frac{1}{\max(\bar{v}^2)}\left[\sum_{i=1}^{n}(x_i s_i - 1 - \bar{v}_i^{-2}\ln x_i s_i)\right], \tag{4}$$

in that the "weights" $\bar{v}_i$ are introduced in the duality gap term instead of the barrier term. The corresponding potential to (4) in terms of v is

$$\tilde{\phi}(v, \bar{v}) = \sum_{i=1}^{n}\frac{\bar{v}_i^2}{\max(\bar{v}^2)}\left[\frac{v_i^2}{\bar{v}_i^2} - 1 - \ln\frac{v_i^2}{\bar{v}_i^2}\right]. \tag{5}$$

Notice that weighting factors $\bar{v}_i^2/\max(\bar{v}^2)$ appear in (5) which are absent from (3). Although the new formulation seems more natural, it suffers from the apparent drawback that it is nonconvex, whereas $\tilde{f}$ in (4) is a strictly convex function of x and s for fixed $\bar{v}$.

Surprisingly, convexity is not a crucial issue here as the two potentials (2) and (4) have the same first-order optimality conditions:

$$\left.\begin{aligned} Ax &= b,\\ A^T y + s &= c,\\ xs &= \bar{v}^2,\\ x, s &\geq 0, \end{aligned}\right\} \tag{6}$$

which are simply the relaxed LP optimality conditions and known to have a unique solution (see e.g. [6]). Moreover, it has already been indicated that the new potential attains its lower bound if $xs = \bar{v}^2$, proving the existence of a unique minimizer of f. In other words, both potential functions have the solution of (6) as a unique minimizer.[2)]

A fixed $\bar{v}$ therefore represents a target which is approached by reducing the potential (2) using Newton's method. Once the potential has been sufficiently reduced, the target can be updated.

3. Reducing the potential

It remains to show that (2) can be successfully minimized by Newton's method. The next theorem shows that a damping parameter $\alpha \leq 1$ can always be found so that the damped Newton step reduces (2) by an absolute constant, determined by the current point (x, s) and the target $\bar{v}$ only.

2) It is interesting to note that the new function does allow a convex reformulation in terms of variables $t_i = x_i s_i/\bar{v}_i^2$, and can be written as $\Psi(t) = \sum_{i=1}^{n}\Psi_i(t_i)$, with $\Psi_i(t_i) = t_i - 1 - \ln t_i$.

Notation: The potential reduction will be given in terms of a function ρ of the distance $\delta(\upsilon, \overline{\upsilon})$ (where no confusion is possible, we will use $\delta := \delta(\upsilon, \overline{\upsilon})$):

$$\rho(\delta) = \delta + \sqrt{1 + \delta^2}.$$

We will also borrow the following notation from Jansen [4]:

$$\begin{aligned} p_x &= \frac{\upsilon \Delta x}{x}, \\ p_s &= \frac{\upsilon \Delta s}{s}, \\ p_\upsilon &= p_x + p_s = \frac{\overline{\upsilon}^2 - \upsilon^2}{\upsilon}, \\ r &= \left\| \left[\frac{p_x}{\upsilon}, \frac{p_s}{\upsilon} \right] \right\|. \end{aligned}$$

Theorem 1. A damped Newton step $(\alpha \Delta x, \alpha \Delta s)$ with damping parameter

$$\alpha = \frac{1}{r} - \frac{\max(\overline{\upsilon})^2}{\| p_\upsilon \|^2 + r \max(\overline{\upsilon})^2} \leq 1 \tag{7}$$

gives a reduction of the potential function (2), bounded by

$$f(x, s, \overline{\upsilon}) - f(x + \alpha \Delta x, s + \alpha \Delta s, \overline{\upsilon}) \geq \frac{2\delta^2 \overline{\varpi}^4}{\rho^2(\delta) + 2\rho(\delta)\delta \overline{\varpi}^2},$$

where $\overline{\varpi} = \max(\overline{\upsilon}) / \min(\overline{\upsilon})$.

Proof. By definition, the reduction of f is given by

$$\begin{aligned} &\Delta f(\alpha) \\ &\quad \equiv f(x, s, \overline{\upsilon}) - f(x + \alpha \Delta x, s + \alpha \Delta s, \overline{\upsilon}) \\ &\quad = -e^T (\alpha^2 \Delta x \Delta s \overline{\upsilon}^{-2} + \alpha x \Delta s \overline{\upsilon}^{-2} + \alpha \Delta x s \overline{\upsilon}^{-2}) + \sum_{i=1}^n \ln\left(1 + \frac{\alpha \Delta x_i}{x_i}\right)\left(1 + \frac{\alpha \Delta s_i}{s_i}\right) \\ &\quad = -e^T\left(\alpha^2 \frac{p_x p_s}{\overline{\upsilon}^2} + \alpha \frac{\upsilon}{\overline{\upsilon}^2}(p_x + p_s)\right) + \sum_{i=1}^n \ln\left(1 + \frac{\alpha (p_x)_i}{\upsilon_i}\right)\left(1 + \frac{\alpha (p_s)_i}{\upsilon_i}\right) \\ &\quad = -e^T\left(\frac{1}{2}\alpha^2 \left[\frac{p_\upsilon}{\overline{\upsilon}}\right]^2 - \frac{1}{2}\alpha^2 \frac{p_x^2 + p_s^2}{\overline{\upsilon}^2} + \alpha \frac{\upsilon}{\overline{\upsilon}^2} p_\upsilon\right) + \sum_{i=1}^n \ln\left(1 + \frac{\alpha (p_x)_i}{\upsilon_i}\right)\left(1 + \frac{\alpha (p_s)_i}{\upsilon_i}\right), \end{aligned}$$

where e denotes the all-one vector and we have used $p_\upsilon = p_x + p_s$. The last term can be bounded by applying the inequality

$$\sum_{i=1}^{n} \ln(1+h_i) \geq \sum_{i=1}^{n} h_i + \|h\| + \ln(1-\|h\|) \quad \text{if } \|h\| \leq 1$$

to the combined vector $h = [\alpha p_x/v, \alpha p_s/v]$. Noting that $\|h\| = \alpha r$ in this case, we obtain

$$\begin{aligned}\Delta f(\alpha) &\geq e^T \left(-\frac{1}{2}\alpha^2 \left[\frac{p_v}{\bar{v}}\right]^2 + \frac{1}{2}\alpha^2 \frac{p_x^2 + p_s^2}{\bar{v}^2} - \alpha \frac{v}{\bar{v}^2} p_v + \alpha \frac{p_v}{v} \right) + \alpha r + \ln(1-\alpha r) \\ &= e^T \left(-\frac{1}{2}\alpha^2 \left[\frac{p_v}{\bar{v}}\right]^2 + \frac{1}{2}\alpha^2 \frac{p_x^2 + p_s^2}{\bar{v}^2} + \alpha \left[\frac{p_v}{\bar{v}}\right]^2 \right) + \alpha r + \ln(1-\alpha r) \\ &= e^T \left(\left(\alpha - \frac{1}{2}\alpha^2\right) \left[\frac{p_v}{\bar{v}}\right]^2 + \frac{1}{2}\alpha^2 \frac{p_x^2 + p_s^2}{\bar{v}^2} \right) + \alpha r + \ln(1-\alpha r).\end{aligned}$$

The factor $(\alpha - \frac{1}{2}\alpha^2)$ is nonnegative, since $\alpha \leq 1$. We therefore have

$$\begin{aligned}\Delta f(\alpha) &\geq \frac{1}{\max(\bar{v}^2)} e^T \left(\left(\alpha - \tfrac{1}{2}\alpha^2\right)[p_v]^2 + \tfrac{1}{2}\alpha^2 (p_x^2 + p_s^2) \right) + \alpha r + \ln(1-\alpha r) \\ &= \alpha \frac{\|p_v\|^2}{\max(\bar{v}^2)} + \alpha r + \ln(1-\alpha r).\end{aligned}$$

The last expression is maximized by

$$\alpha^* = \frac{1}{r} - \frac{\max(\bar{v})^2}{\|p_v\|^2 + r\max(\bar{v}^2)}, \tag{8}$$

which corresponds to

$$\Delta f(\alpha^*) \geq \frac{\|p_v\|^2}{r\max(\bar{v}^2)} - \ln\left(1 + \frac{\|p_v\|^2}{r\max(\bar{v}^2)}\right). \tag{9}$$

The lower bound (9) on $\Delta f(\alpha^*)$ is obviously nonnegative but must be bounded away from zero. To accomplish this, note that expression (9) increases monotonically with $\|p_v\|^2/(r\max(\bar{v}^2))$. We can therefore replace this quantity by a smaller value. Jansen [4] shows that

$$\frac{\|p_v\|^2}{r\max(\bar{v}^2)} \geq \frac{2\delta\bar{\omega}^2}{\rho(\delta)},$$

where $\bar{\omega} = \max(\bar{v})/\min(\bar{v})$. It follows that

$$\Delta f(\alpha^*) \geq \frac{2\delta\bar{\omega}^2}{\rho(\delta)} - \ln\left(1 + \frac{2\delta\bar{\omega}^2}{\rho(\delta)}\right).$$

Using the inequality

$$x - \ln(1+x) \geq \frac{x^2}{2(x+1)},$$

we arrive at the bound in the theorem statement. □

To fix our ideas, we choose a threshold value to decide when the current iterate υ is "close enough" to the current target $\overline{\upsilon}$. Following Jansen, we use $\tau = \frac{1}{4}$ as the threshold value. As long as $\delta(\upsilon, \overline{\upsilon}) \geq \frac{1}{4}$, we perform damped Newton steps with respect to $\overline{\upsilon}$ with the following guaranteed reduction of f each time:

Corollary 2. If $\delta(\upsilon, \overline{\upsilon}) \geq \frac{1}{4}$, then

$$\Delta f \geq \frac{\overline{\omega}^4}{14 + 6\overline{\omega}^2}.$$

The actual reduction obtained from a line search is of course much larger in general.

Once the proximity condition is satisfied, an upper bound on the potential is also known:

Lemma 3. If $\delta(\upsilon, \overline{\upsilon}) \leq \frac{1}{4}$, then $\phi(\upsilon, \overline{\upsilon}) \leq \frac{2}{5}$.

Proof. The potential ϕ in (3) can be written as

$$\begin{aligned}\phi(\upsilon, \overline{\upsilon}) &= \sum_{i=1}^{n} \left(\frac{\upsilon_i^2}{\overline{\upsilon}_i^2} - 1 \right) - \sum_{i=1}^{n} \ln \frac{\upsilon_i^2}{\overline{\upsilon}_i^2} \\ &= \sum_{i=1}^{n} h_i - \sum_{i=1}^{n} \ln(1 + h_i),\end{aligned}$$

where $h_i = (\upsilon_i^2 / \overline{\upsilon}_i^2) - 1$. Since

$$\|h\| = \left\| \frac{\upsilon}{\overline{\upsilon}^2} \, \frac{e}{\upsilon} (\upsilon^2 - \overline{\upsilon}^2) \right\| \leq \left\| \frac{\upsilon}{\overline{\upsilon}^2} \right\|_\infty \| p_\upsilon \| \leq \rho(\delta) \| \overline{\upsilon}^{-1} \|_\infty 2 \min(\overline{\upsilon}) \delta = 2\delta\rho(\delta) < \tfrac{13}{20} < 1,$$

if $\delta \leq \frac{1}{4}$, one can use the inequality

$$\sum_{i=1}^{n} h_i - \sum_{i=1}^{n} \ln(1 + h_i) \leq -\|h\| - \ln(1 - \|h\|),$$

which holds if $\|h\| < 1$, to obtain

$$\phi(\upsilon, \overline{\upsilon}) \leq -2\delta\rho(\delta) - \ln(1 - 2\delta\rho(\delta)).$$

Since $\delta(\upsilon, \overline{\upsilon}) \leq \frac{1}{4}$ and consequently $\delta\rho(\delta) < \frac{13}{40}$, we have $\phi(\upsilon, \overline{\upsilon}) \leq \frac{2}{5}$. □

All the tools necesserary to control the Newton process have now been developed, and we turn to the analysis of target updates.

4. Analysis of a general target update

Once the current iterate v is close enough to the target $\overline{v}$, i.e. $\delta(v, \overline{v}) \le \frac{1}{4}$, the target can be updated to $\overline{v}^+$. The new potential $\phi(v, \overline{v}^+)$ can be bounded from above as follows:

Lemma 4. Given a current iterate v, current target $\overline{v}$, and target update $\overline{v}^+$, it holds that

$$\phi(v, \overline{v}^+) \le \phi(\overline{v}, \overline{v}^+) + \max\left(\frac{\overline{v}^2}{(\overline{v}^+)^2}\right)\phi(v, \overline{v}) + 4\delta\rho(\delta)\sqrt{n}\,\max\left(\left|\frac{\overline{v}^2}{(\overline{v}^+)^2} - e\right|\right).$$

Proof. The potential after the target update can be written as

$$\begin{aligned}
\phi(v, \overline{v}^+) &= \sum_{i=1}^n \left(\frac{v_i^2}{\overline{v}_i^2}\,\frac{\overline{v}_i^2}{(\overline{v}_i^+)^2}\right) - \sum_{i=1}^n \ln\frac{\overline{v}_i^2}{(\overline{v}_i^+)^2} - \sum_{i=1}^n \ln\frac{v_i^2}{\overline{v}_i^2} - n \\
&= \phi(\overline{v}, \overline{v}^+) + \sum_{i=1}^n \left(\frac{v_i^2}{\overline{v}_i^2} - 1\right)\frac{\overline{v}_i^2}{(\overline{v}_i^+)^2} - \sum_{i=1}^n \ln\frac{v_i^2}{\overline{v}_i^2} \\
&\le \phi(\overline{v}, \overline{v}^+) + \max\left(\frac{\overline{v}^2}{(\overline{v}^+)^2}\right)\phi(v, \overline{v}) + \sum_{i=1}^n \left(\frac{\overline{v}_i^2}{(\overline{v}_i^+)^2} - 1\right)\ln\frac{v_i^2}{\overline{v}_i^2}.
\end{aligned}$$

In [4, theorem 4.3.6], it is proved that the last term is bounded by

$$\sum_{i=1}^n \left(\frac{\overline{v}_i^2}{(\overline{v}_i^+)^2} - 1\right)\ln\frac{v_i^2}{\overline{v}_i^2} \le 4\delta\rho(\delta)\sqrt{n}\,\max\left(\left|\frac{\overline{v}^2}{(\overline{v}^+)^2} - e\right|\right).$$

Substitution of this bound completes the proof. □

By combining corollary 2 and lemma 4, the following result is obtained.

Lemma 5. If the current iterate $v^{(j)}$ satisfies $\delta(v^{(j)}, \overline{v}) \le \frac{1}{4}$, and the target $\overline{v}$ is updated to $\overline{v}^+$, then at most

$$\left\lceil \frac{14 + 6\varpi^2}{\varpi^4}\left[\phi(\overline{v}, \overline{v}^+) + \tfrac{2}{5}\max\left(\frac{\overline{v}^2}{(\overline{v}^+)^2}\right) + \tfrac{13}{10}\sqrt{n}\,\max\left(\left|\frac{\overline{v}^2}{(\overline{v}^+)^2} - e\right|\right)\right]\right\rceil$$

damped Newton steps with respect to $\overline{v}^+$ are required to obtain an iterate $v^{(j+1)}$ satisfying $\delta(v^{(j+1)}, \overline{v}^+) \le \frac{1}{4}$.

5. Complexity analysis for Dikin-type target updates

All that remains is to choose a target updating scheme. Consider for example the Dikin-type updates introduced by Jansen et al. [5–7]:

$$\overline{v}^{+} = \overline{v}\left(e - \frac{\theta}{\max(\overline{v}^{2\nu})}\,\overline{v}^{2\nu}\right), \tag{10}$$

with $0 < \theta < 1/(2\nu + 1)$. Note that $\nu = 0$ corresponds to weighted path-following methods. Furthermore, an initial choice $\overline{v}^{(0)} = \mu e$ for some fixed $\mu > 0$ leads to a central path-following algorithm.

We can bound the number of Newton steps necessary to approximate a new target $\overline{v}^{+}$ given $\delta(v, \overline{v}) \le \frac{1}{4}$ by providing bounds for each of the terms in lemma 5.

Lemma 6. Let $\overline{v}^{+}$ be a new target obtained by updating the old target $\overline{v}$ via (10). We then have the bounds

$$\phi(\overline{v}, \overline{v}^{+}) \le \frac{3n\theta^2}{(1-\theta)^2}, \quad \max\left(\frac{\overline{v}^2}{(\overline{v}^{+})^2}\right) \le \frac{1}{(1-\theta)^2}, \quad \max\left(\left|\frac{\overline{v}^2}{(\overline{v}^{+})^2} - e\right|\right) \le \frac{\theta(2-\theta)}{(1-\theta)^2}.$$

Proof. The last two inequalities follow from the observation

$$\frac{\overline{v}_i^2}{(\overline{v}_i^{+})^2} = \frac{1}{(1 - \theta\overline{v}_i^{2\nu}/\max(\overline{v}^{2\nu}))^2} \le \frac{1}{(1-\theta)^2}.$$

The bound on $\phi(\overline{v}, \overline{v}^{+})$ is obtained as follows:

$$\begin{aligned}
\phi(\overline{v}, \overline{v}^{+}) &\equiv \sum_{i=1}^{n} \frac{\overline{v}_i^2}{(\overline{v}_i^{+})^2} - \sum_{i=1}^{n}\left(\ln\frac{\overline{v}_i^2}{(\overline{v}_i^{+})^2} - 1\right) \\
&= \sum_{i=1}^{n} \frac{1}{(1 - \theta\overline{v}_i^{2\nu}/\max(\overline{v}^{2\nu}))^2} - \sum_{i=1}^{n}\left(\ln\frac{1}{(1 - \theta\overline{v}_i^{2\nu}/\max(\overline{v}^{2\nu})^2} - 1\right) \\
&= \sum_{i=1}^{n} \frac{1}{(1 - \theta\overline{v}_i^{2\nu}/\max(\overline{v}^{2\nu}))^2}\left[1 + 2\left(1 - \frac{\theta\overline{v}_i^{2\nu}}{\max(\overline{v}^{2\nu})}\right)^2\left(\ln\left(1 - \frac{\theta\overline{v}_i^{2\nu}}{\max(\overline{v}^{2\nu})}\right)\right.\right. \\
&\qquad \left.\left. - \left(1 - \frac{\theta\overline{v}_i^{2\nu}}{\max(\overline{v}^{2\nu})}\right)^2\right)\right].
\end{aligned}$$

Using $\ln(1-x) \le -x$ if $x < 1$ and simplifying, we have

$$\phi(\overline{v}, \overline{v}^+) \leq \sum_{i=1}^{n} \frac{1}{(1-\theta\overline{v}_i^{2\nu}/\max(\overline{v}^{2\nu}))^2}\left[2\left(1-\frac{\theta\overline{v}_i^{2\nu}}{\max(\overline{v}^{2\nu})}\right)^2\left(\frac{-\theta\overline{v}_i^{2\nu}}{\max(\overline{v}^{2\nu})}\right)^2 + 2\frac{\theta\overline{v}_i^{2\nu}}{\max(\overline{v}^{2\nu})} - \frac{\theta^2\overline{v}_i^{4\nu}}{\max(\overline{v}^{4\nu})}\right]$$

$$\leq \frac{1}{(1-\theta)^2}\sum_{i=1}^{n}\left(3\frac{\theta^2\overline{v}_i^{4\nu}}{\max(\overline{v}^{4\nu})} - 2\frac{\theta^3\overline{v}_i^{6\nu}}{\max(\overline{v}^{6\nu})}\right) \leq \frac{3n\theta^2}{(1-\theta)^2}.$$

□

We now have a bound on how many damped Newton steps are required to reach the proximity of a new target:

Corollary 7. Assume that $\delta(v, \overline{v}) \leq \frac{1}{4}$. If the target is updated to $\overline{v}^+$ using the target updating scheme (10), then at most

$$O\left(\frac{1}{(\overline{\omega}^+)^4}(n\theta^2 + \sqrt{n}\theta)\right)$$

damped Newton steps are needed to approximate $\overline{v}^+$, where $\overline{\omega}^+ = \max(\overline{v}^+)/\min(\overline{v}^+)$.

The last question is how many target updates are required to obtain an ε-approximate solution. It is simple to prove the following (see [4]):

Lemma 8. Let a primal-dual starting pair $(x^{(0)}, s^{(0)})$ be given. Choose the first target as $\overline{v}^{(0)} = v^{(0)}$. After at most

$$O\left(\frac{1}{\theta\omega_0^{2\nu}}\ln\frac{(x^{(0)})^T s^{(0)}}{\varepsilon}\right)$$

target updates using (10), the algorithm terminates with a primal-dual pair (x^*, s^*) such that $(x^*)^T s^* \leq \varepsilon$.

Combining these results, we obtain the complexity bound for the complete algorithm:

Theorem 9. The target-following algorithm requires at most

$$O\left(\frac{n\theta + \sqrt{n}}{\omega_0^{2\nu+4}}\ln\frac{(x^{(0)})^T s^{(0)}}{\varepsilon}\right)$$

damped Newton steps for convergence.

A large target update with $\theta = O(1)$ therefore requires fewer than $O(n/\omega_0^{2\nu+4})$ Newton steps, whereas medium step methods with $\theta = O(1/\sqrt{n})$ require fewer than $O(\sqrt{n}/\omega_0^{2\nu+4})$ steps.

These complexity bounds are the same as those obtained by using the standard convex potential function. We conclude that the nonconvex potential (3) is a proper alternative to the usual convex logarithmic barrier potential.

6. Further work

It has already been mentioned that the new potential function (2) has an extension to the semi-definite programming (SDP) case. The recent revival of interest in SDP started more or less with the work of Alizadeh [1] and Nesterov and Nemirovskii [14]; an excellent review on developments and applications up to 1995 is given by Vandenberghe and Boyd in [15]. One reason for the recent interest in SDP is that most interior-point methods for LP can be extended to SDP. This is presently an active research area, as can be seen from the number of recent publications (see e.g. [10,12,16]).

The general semi-definite problem can be formulated as

$$\begin{aligned} &\min \operatorname{Tr}(CX) \\ &\operatorname{Tr}(A_i X) = b_i \quad i = 1, \ldots, m, \\ &X \succeq 0, \end{aligned}$$

where the A_i's and C are symmetric matrices, "$\succeq$" denotes positive semi-definiteness, and "Tr" denotes the trace of a matrix. The associated dual problem is

$$\begin{aligned} &\max b^T y \\ &\textstyle\sum_{i=1}^m y_i A_i + S = C, \\ &S \succeq 0, \end{aligned}$$

and at optimality, one has $XS = 0$. Note that the potential function

$$f(X, S, \overline{V}) = \operatorname{Tr}(XS\overline{V}^{-2}) - \ln\det\left(\tfrac{1}{2}(XS + SX)\overline{V}^{-2}\right) - n \tag{11}$$

is a natural extension of the new LP potential (2) to the semi-definite case, where $\overline{V}$ is a symmetric positive definite "target matrix". (Note that $\frac{1}{2}(XS + SX)$ is the symmetric part of XS.)

This potential function has a unique minimizer over all primal-dual feasible pairs (X, S) with $XS + SX$ symmetric positive definite. This follows from a recent paper by Monteiro and Pang [13], where it is shown that for each symmetric positive definite $\overline{V}$, there exists a unique primal-dual feasible pair satisfying $\frac{1}{2}(XS + SX) = \overline{V}$.

To the best of our knowledge, the function defined in (11) is the first weighted potential function for semi-definite programming. Extension of the analysis in the previous sections would therefore broaden the target-following framework to semi-definite programming. This is the subject of further research.

References

[1] F. Alizadeh, Combinatorial optimization with interior point methods and semi-definite matrices, Ph.D. Thesis, University of Minnesota, Minneapolis, MN, 1991.

[2] C.C. Gonzaga, Large steps path-following methods for linear programming, Part I: Barrier function method, SIAM Journal on Optimization 1(1991)268 – 279.

[3] D. den Hertog, *Interior Point Approach to Linear, Quadratic and Convex Programming, Algorithms and Complexity*, Kluwer, Dordrecht, The Netherlands, 1994.

[4] B. Jansen, Interior point techniques in optimization, *Complementarity, Sensitivity and Algorithms*, Kluwer, Dordrecht, The Netherlands, 1997.

[5] B. Jansen, C. Roos and T. Terlaky, A polynomial Dikin-type primal-dual algorithm for linear programming, Math. of Operations Research 21(1996)341 – 353.

[6] B. Jansen, C. Roos, T. Terlaky and J.-Ph. Vial, Primal-dual target-following algorithms for linear programming, Annals of Operations Research 62(1996)197 – 231.

[7] B. Jansen, C. Roos, T. Terlaky and J.-Ph. Vial, Long-step primal-dual target-following algorithms for linear programming, Mathematical Methods of Operations Research 44(1996)11 – 30.

[8] B. Jansen, C. Roos, T. Terlaky and J.-Ph. Vial, Primal-dual algorithms for linear programming based on the logarithmic barrier method, Journal of Optimization Theory and Applications 83(1994) 1 – 26.

[9] M. Kojima, S. Mizuno and A. Yoshise, A primal-dual interior point algorithm for linear programming, in: *Progress in Mathematical Programming: Interior Point and Related Methods*, ed. N. Megiddo, Springer, New York, 1989, pp. 29 – 47.

[10] M. Kojima, M. Shida and S. Shindoh, Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs, Mathematical Programming 80(1998)129 – 160.

[11] I.J. Lustig, R.E. Marsten and D.F. Shanno, Interior point methods: Computational state of the art, ORSA Journal on Computing 6(1994)1 – 15.

[12] R.D.C. Monteiro, Primal-dual algorithms for semidefinite programming, SIAM J. Optimization 7(1997)663 – 678.

[13] R.D.C. Monteiro and J.-S. Pang, On two interior point mappings for nonlinear semidefinite complementarity problems, Mathematics of Operations Research 23(1998)39 – 60.

[14] Yu. Nesterov and A.S. Nemirovskii, Interior point polynomial algorithms in convex programming, *SIAM Studies in Applied Mathematics*, Vol. 13, SIAM, Philadelphia, PA, 1994.

[15] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Review, 38(1996)49 – 95.

[16] Y. Zhang, On extending some primal-dual interior point algorithms from linear programming to semi-definite programming, to appear in SIAM J. Optimization (1998).